CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	3D ranging and tracking using lensless smart sensors
Authors	Abraham, Lizy;Urru, Andrea;Wilk, Mariusz P.;Tedesco, Salvatore;O'Flynn, Brendan
Publication date	2017-03
Original Citation	Abraham, L., Urru, A., Wilk, M., Tedesco, S. and O'Flynn, B. (2017) '3D Ranging and Tracking Using Lensless Smart Sensors', 11th Smart Systems Integration, SSI 2017: International Conference and Exhibition on Integration Issues of Miniaturized Systems, Cork, Ireland, 8-9 March. isbn 978-3-95735-057-2
Type of publication	Conference item
Link to publisher's version	https://www.smart-systems-integration.org/public
Download date	2023-05-05 00:55:00
Item downloaded from	http://hdl.handle.net/10468/5626

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

3D Ranging and Tracking Using Lensless Smart Sensors

Lizy Abraham, Andrea Urru, Mariusz Wilk, Salvatore Tedesco, Brendan O'Flynn
Tyndall National Institute, University College Cork, Cork, Ireland;

1. Introduction

Target tracking has a wide range of applications in Internet of Things (IoT), such as smart city sensors, indoor tracking, and gesture recognition. Several studies have been conducted in this area. Most of the published works either use vision sensors or inertial sensors for motion analysis and gesture recognition [1, 2]. Recent works use a combination of depth sensors and inertial sensors for 3D ranging and tracking [3, 4]. This often requires complex hardware and the use of complex embedded algorithms. Stereo cameras or Kinect depth sensors used for high precision ranging are instead expensive and not easy to use. The aim of this work is to track in 3D a hand fitted with a series of precisely positioned IR LEDs using a novel Lensless Smart Sensor (LSS) developed by Rambus, Inc. [5, 6]. In the adopted device, the lens used in conventional cameras is replaced by low-cost ultra-miniaturized diffraction optics attached directly to the image sensor array. The unique diffraction pattern enables more precise position tracking than possible with a lens by capturing more information about the scene.

2. Methodology

For the 3D ranging, two LSS sensors (Sen_R and Sen_L) were adopted and, by combining the raw images obtained from both of them, it is possible to calculate the position of each LED in the 3D space with respect to the sensor plane (e.g. the plane identified by the axes X and Y, as shown in Fig.1).
The stereoscopic Field-of-View (FoV) is defined on the basis of the following parameters:

(i) the longitudinal distance on the Z-axis between the LED and the sensor plane;
(ii) the distance between Sen_R and Sen_L measured on the central points of the sensors, called baseline;

(iii) the combined Angle-of-View (AoV).

The light source adopted in the set-up is a low-cost infrared LED and the sensors are equipped with an infrared filter to mitigate the effects due to ambient light.

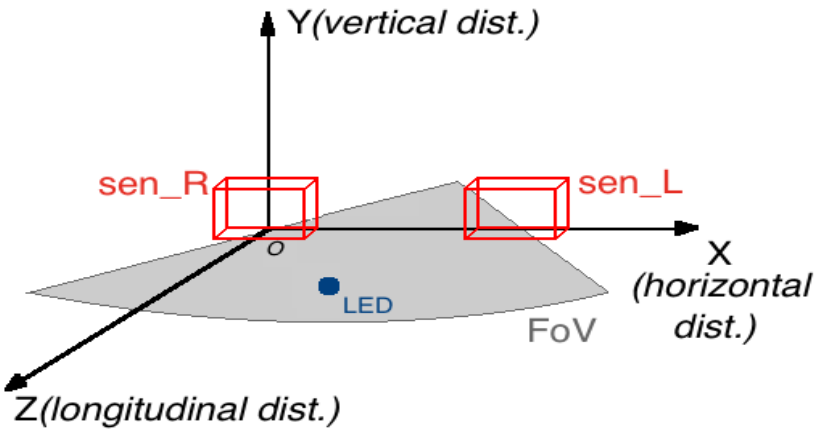

Fig. 1. System Overview

2.1 Image Reconstruction

The latest version of LSS has two apertures and creates a dual spiral binary diffractive pattern on the image sensor for a single point source of light. In order to reconstruct the light points created by the LEDs in the image frame accurately, an optimal Point Spread Function (PSF) is generated with a low noise-level. To this purpose, the reference PSF is acquired by subtracting two reference frames both estimated averaging 1000 images of the LED located at 1 m from the sensor plane in On and Off conditions, respectively, as shown in [6, 7]. Then, the regularized inverse of the Fourier transformation of the reference PSF, also called kernel, is calculated. Finally, the final reconstructed image is obtained through the inverse Fourier transform of the product between the kernel and the Fourier-transformed raw image. The two spirals on the dual gratings are not identical to each other; indeed, one is rotated 22 deg with respect to the other so that one spiral could cover potential deficiencies present in the other sensor, and *vice versa*. Therefore, it is required to use the reconstruction result for each spiral separately and then add them together. This also allows to effectively double the amount of light captured, thus reducing noise effects. Both sides of the image are reconstructed with their

respective kernels and then added together. Figure 2 shows an example of the reference PSF for dual spiral diffraction optics and the reconstructed point.

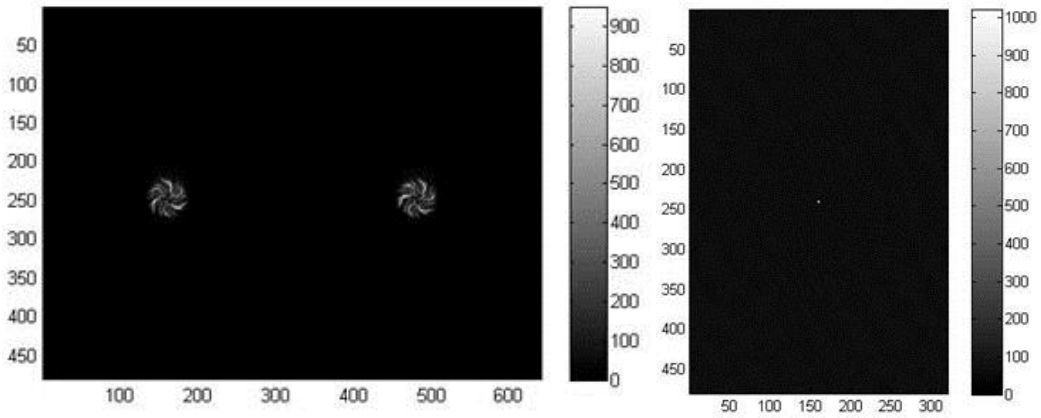

Fig. 2: Reference PSF (left) and the reconstructed point (right)

2.2 3D Ranging

After accurately reconstructing the image frame of a LED, it is possible to estimate the 3D ranging between the origin of the reference system, on the focal point (FP) of the right LSS, and the LED. As shown in Fig. 3, *b* is the defined baseline, considered fixed during the test, and is equal to 15cm, *A* and *B* are the horizontal distances respectively between the right and left FPs and the points where the photons hit the surface of the LSS, while *C* and *D* are related vertical distances.

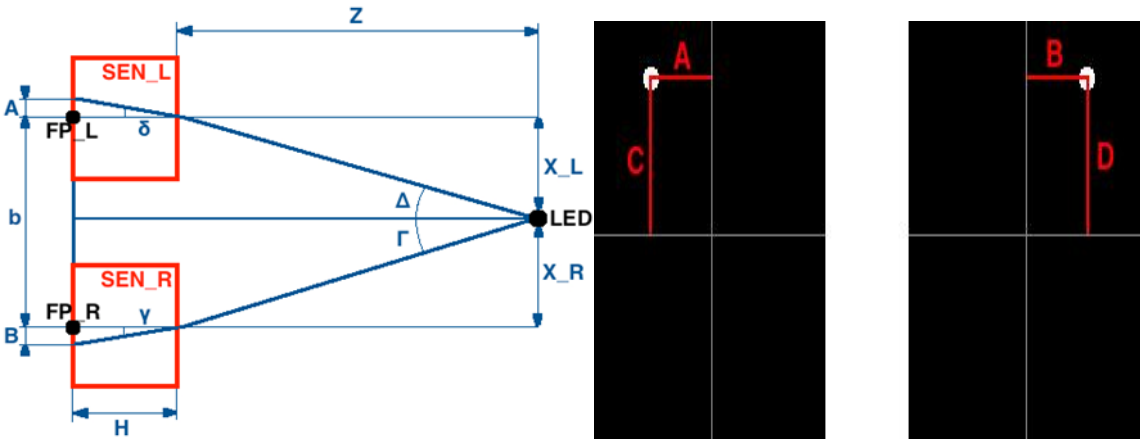

Fig. 3: 3D Ranging Geometry

The longitudinal distance Z can be calculated using trigonometry and Snell's law of refraction [5, 8] and, consequently, the lateral and vertical displacements (along X and Y, respectively) of the LED with respect to the origin can be derived according to the following equations. Starting from the law of cosines and basic trigonometry, it is obtained that

$$b^2 = Z^2 + X_L^2 + Z^2 + X_R^2 - 2\sqrt{Z^2 + X_L^2}\sqrt{Z^2 + X_R^2}\cos(\Delta + \Gamma) \quad (1)$$

$$\tan(\Delta) = \frac{X_L}{Z} \quad (2)$$

$$\tan(\Gamma) = \frac{X_R}{Z} \quad (3)$$

H is the distance from the end of the LSS to its FP and is a technological parameter based on the materials stack-up. It is known *a priori* and assumed equal to 465 microns, and allows the calculation of the angles δ and γ as below in (4)-(5).

$$\tan(\delta) = \frac{A}{H} \quad (4)$$

$$\tan(\gamma) = \frac{B}{H} \quad (5)$$

According to Snell's law,

$$\sin(\Delta) = n\sin(\delta) \quad (6)$$

$$\sin(\Gamma) = n\sin(\gamma) \quad (7)$$

where n is the total refractive index of the sensor and it is assumed equal to 1.58. Therefore, substituting equations (4)-(5) in (6)-(7), it is possible to obtain Δ and Γ. Substituting X_L and X_R from equations (2)-(3) in (1), Z can be then calculated as:

$$Z = \frac{b}{\sqrt{\left[2 + \tan^2(\Delta) + \tan^2(\Gamma) - 2\sqrt{1 + \tan^2(\Delta)}\sqrt{1 + \tan^2(\Gamma)}\cos(\Delta + \Gamma)\right]}} \quad (8)$$

Thus, X_L and X_R can be calculated from Z using (2)-(3). Similarly, Y is obtained repeating the same process after replacing A and B with the measured vertical distances C and D.

2.3 Distortion Compensation

Finally, the overall accuracy of the ranging algorithm is improved by compensating the error due to shot noise in the estimation of the angles Δ and Γ using a polynomial fit function. The distortion compensation is modelled by collecting a series of images of the LED moved on a line parallel to the sensor plane. The longitudinal distance is 40cm from sensor plane, and data are collected for a total of 61 positions, spaced 1 cm, along the X-axis in order to cover the whole AoV. For each point, 5 raw image frames were acquired by both sensors. A polynomial fit function is then obtained by calculating the coefficients of the fitting curve which minimizes the root mean square (RMS) error.

Finally, the ranging algorithm combined for both LSS simultaneously makes a real-time single point tracking possible.

2.4 Multi-Point Live Tracking

By adopting a series of infrared LEDs on a hand model, the live tracking of the body part is obtained. In order to minimize both occlusions and the number of LEDs adopted, yet allowing the recognition of the two sides of the hand and the vision, at the same time, of at least 3 lights by the two LSS, a careful placement has been studied which results in attaching one LED on the front palm, three on the back palm, one on each side of the wrist, and a final one on the frontal side of the thumb. By running the above described ranging algorithm steadily for each LED, it is possible to establish location and orientation of the hand in real-time.

Figure 4 visualizes the hand tracking from the accurate positioning of LEDs.

3. Results and Conclusions

A novel method for 3D ranging and tracking using the novel LSS developed by Rambus, Inc. is described. The algorithms are implemented in Matlab and are capable of dynamically recognizing the different sides of the hand and computing its

positioning even if equipped with minimum number of LEDs.

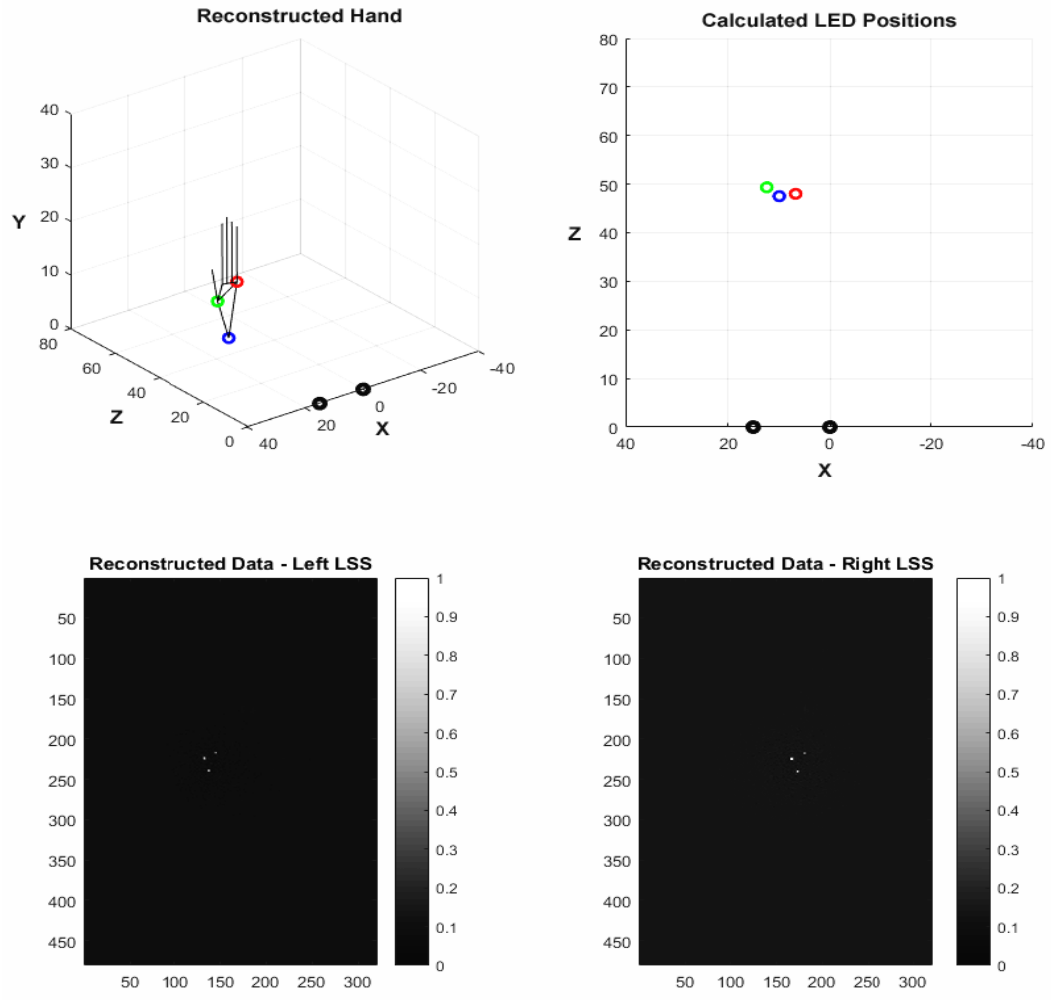

Fig. 4: Hand Tracking Via Multiple LEDs

Figure 5-6 show the positive effects of the distortion compensation on the overall performance of the algorithm. The case analyzed is at longitudinal distance of 40 cm. Table 1 illustrates the accuracy of the single point live tracking algorithm in terms of RMS error with and without compensation function. It is evident that the compensation reduces the error to less than 1 cm along the Z-axis and 0.34 cm along the X-axis, resulting in a significant improvement. Table 2 shows the RMS error for three different cases: along the centre of the baseline, and on the external sides of the system, with the angle α, defining the AoV, equal to ±40 deg. For all three

cases, the total RMS error is then calculated by varying the longitudinal distance from 40 to 100 cm with 10 cm spacing. The RMS error obtained along the Z-axis, within the AoV of 80 deg given by placing the LSS 15cm apart, is lower than 2.5 cm in all three cases. Moreover, along the central axis, the error is limited to 0.5 cm.

The study proved that the LSS sensor developed by Rambus, Inc. can be adopted for live tracking of multiple light points. The overall system does not require complex hardware or additional equipment on the tracked segments. This innovative technology has a range of potential applications in IoT (e.g. virtual reality) and in the future it will be integrated into a wearable system for hand tracking and gesture recognition.

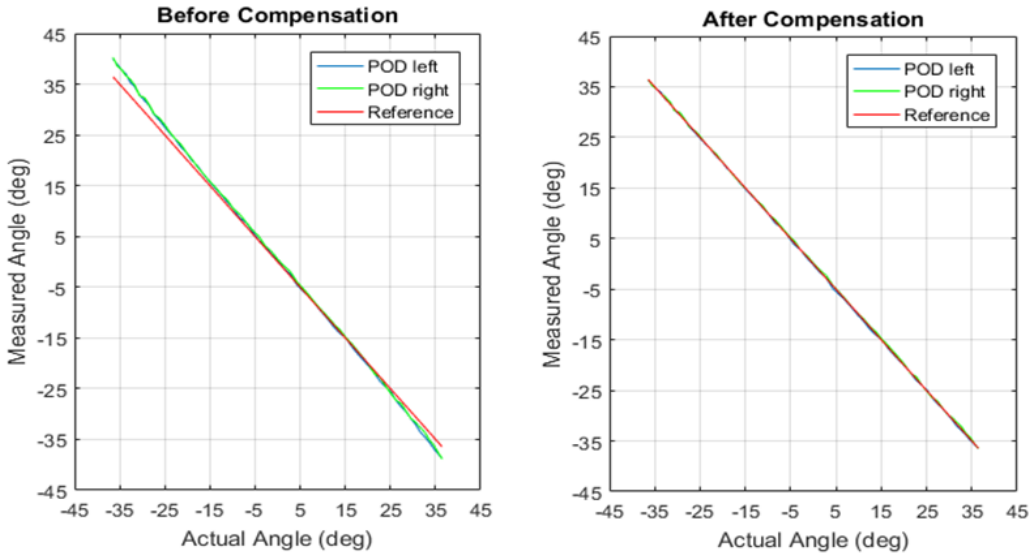

Fig. 5: Angle Correction

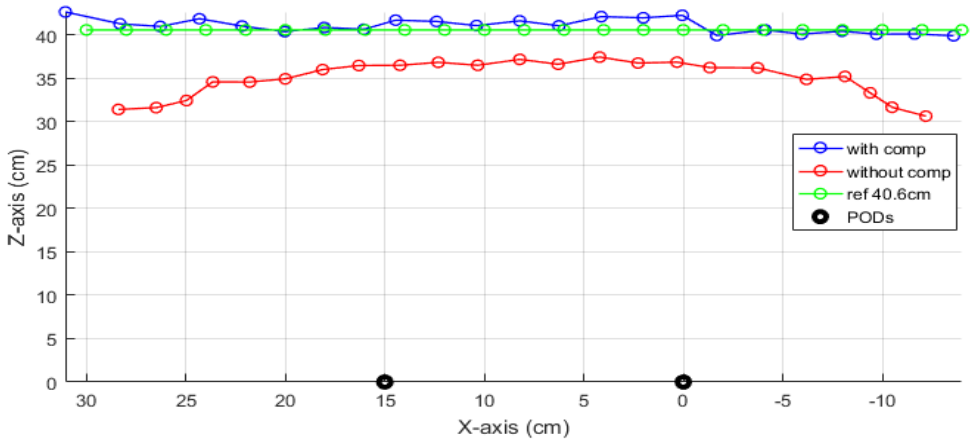

Fig. 6: Distortion Compensation (at Z = 40 cm)

	RMSE – Z (cm)	RMSE – X (cm)
with Compensation	0.9186	0.3368
without Compensation	5.9526	0.7612

Table 1. Accuracy Analysis (at Z = 40 cm)

	α = 0 deg	α = -40 deg	α = +40 deg
RMSE along Z-Axis (cm)	0.29	2.13	2.47

Table 2. RMS error for different angles (Z between 40 and 100 cm)

References

[1] P. Plawiak, *et al.*, “Hand Body Language Gesture Recognition Based on Signals from Specialized Glove and Machine Learning Algorithms,” IEEE Trans. Ind. Inf., 12(3), 2016.

[2] Z. Ren, *et al.*, “Robust Part-Based Hand Gesture Recognition Using Kinect Sensor,” IEEE Trans. Multim., 15(5), 1110–1120, 2013.

[3] Y. Ren, *et al.*, “Hand Gesture Recognition with Multi-Scale Weighted Histogram of Contour Direction (MSWHCD) Normalization for Wearable Applications,” IEEE Trans. Circ. Sys. Video Tech., 2016.

[4] K. Liu, *et al.*, “Fusion of Inertial and Depth Sensor Data for Robust Hand Gesture Recognition,” IEEE Sens. J., 14(6), 1898-1903, 2014.

[5] P. Gill, *et al.*, “Lensless Smart Sensors: Optical and Thermal Sensing for the Internet of Things,” IEEE Symp. VLSI Circ., 2016.

[6] E. Erickson, *et al.*, “Miniature Lensless Computational Infrared Imager,” Elect. Imag., Im. Sens. Imag. Sys. 2016, 1-4 (4).

[7] D.G. Stork, *et al.*, “Optical, Mathematical, and Computational Foundations of Lensless Ultra-Miniature Diffractive Imagers and Sensors”, International Journal on Advances in Systems and Measurements, 7(3-4), 201-208, 2014.

[8] J. Mrovlje, *et al.*, “Distance Measuring based on Stereoscopic Pictures”, 9th International PhD Workshop on Systems and Control, 1-6, 2008.

View publication stats